

Response experiments for nonlinear systems with application to reaction kinetics and genetics

Marcel O. Vlad^{*†}, Adam Arkin[‡], and John Ross^{*§}

^{*}Department of Chemistry, Stanford University, Stanford, CA 94305-5080; [†]Department of Bioengineering and Chemistry, University of California, Berkeley, CA 94720; and [‡]Center of Mathematical Statistics, Casa Academiei Romane, Calea 13 Septembrie 13, Buharest 5, Romania

Contributed by John Ross, March 26, 2004

A unified type of response experiment is suggested for complex systems made up of individual species (atoms, molecules, quasiparticles, biological organisms, etc.). We make the following assumptions: (i) some of the species may exist in two forms, labeled and unlabeled, respectively; (ii) the kinetic and transport properties of the labeled and unlabeled species are the same, respectively (neutrality assumption); (iii) the experiment preserves the total input and output fluxes; only the fractions of the labeled compounds in the input and output fluxes are varied. Under these circumstances a linear integral superposition law connects the fractions of labeled species in the input and output fluxes. This linear superposition law is valid for homogeneous and inhomogeneous systems and for systems with intrinsic (hidden) state variables; it arises from the neutrality condition and holds even though the underlying dynamics of the process may be highly nonlinear. Because this response law does not involve the linearization of the evolution equations it has great potential for the analysis of complex physical, chemical, and biological systems. We compare our approach with the linearization techniques used in biochemistry and genetics. We consider a simple reaction network involving replication, transformation, and disappearance steps and study the influence of experimental (measurement) and linearization errors on the evaluated values of rate coefficients. We show that the method involving the linearization of the kinetic equations leads to unpredictable results; because of the interference between measurement and linearization errors, either error compensation or error amplification occurs. Although our approach does not eliminate the effects of measurement errors, it leads to more consistent results. For a broad range of input fractions no error amplification or compensation occurs, and the error range for the rate coefficients is about the same as the error range of the measurements.

The study of large chemical and genetic networks presents several problems. The measured quantities are concentrations of various chemical species in chemical and biological networks, proteins expressed by genes, and mRNA in genomic and proteomic networks (1). The measurements are made frequently, but not always, from a few to many times. The precision of these measurements varies from at best one, or a few percent, to merely qualitative estimates of 30–50%.

The challenge of the analysis of such measurements, in particular, the deduction of the structure of the network, has been attempted in several ways. We mention a few representative methods: correlation metric construction (2); experiments of pulses of arbitrary magnitude (3); the method of singular-value decomposition, which has been used in two separate ways for the analysis of genetic networks (4, 5); the use of first-order derivatives (Jacobians) (6, 7) and finally, the application of Boolean networks, which have been used as a model for the so-called reverse engineering problem (ref. 8 and references therein).

Response experiments of the type mentioned play important roles not only in chemistry and molecular biology but also in other areas of science and technology, studying materials sub-

jected to variable external electric or magnetic fields (9), the impact of spreading pollutants on the environment (10), and the influence of the occurrence of certain mutations in the past on the current geographic distributions of gene frequencies (11). The response technique consists of inducing controlled perturbations of certain input parameters and of recording the time variation of some output variables that describe the state of the system. The objective of the research is to extract information about the mechanism of the processes occurring within the system. A typical example is the use of tracer experiments for identifying reaction mechanisms in chemical kinetics or molecular biology (ref. 12 and references therein). A similar approach involves perturbations that are naturally occurring, for example, the study of the influence of the appearance of a mutation in the past on the current geographic distribution of gene frequencies in humans (11, 13).

The use of small perturbations leads to important simplifications. If the functional dependence between the excitations and responses is analytic, then for small perturbations it can be approximated by a linear functional dependence, which expresses the contribution of different values of the excitation variables $x_u(t')$, at different times $-\infty < t' \leq t$, on the current values $y_u(t)$ of the response variables at the current time t :

$$y_u(t) = \sum_{u'} \int_{-\infty}^t \mathfrak{E}_{uu'}(t; t') x_{u'}(t') dt', \quad [1]$$

where $\mathfrak{E}_{uu'}(t; t')$ are two-time susceptibility functions that bear the mechanistic information. The important fact is that the linear response law (Eq. 1) is independent of the detailed form of the evolution equations; for small perturbations many different evolution equations may lead to the same linear response law (Eq. 1). In some cases, the linear response law (Eq. 1) can be further simplified, for example, if the susceptibility functions can be represented by the linear combination of a finite number of exponentials, then the dynamics of the system can be represented by a system of linear differential equations. Linear differential equations have been used in approximate studies of gene expression (5).

It is desirable to identify observations or design experiments for which the validity range of a linear response law is known. We have recently introduced two related approaches that lead to linear response laws for systems with nonlinear dynamics. The first approach is based on the use of labeled compounds for identifying reaction mechanisms in chemical kinetics and involves discrete state variables in continuous time (ref. 12 and references therein). The second approach concerns studies of geographic distributions of gene frequencies in humans and uses a continuous time and space description (13, 14). In both cases we have shown that linear response laws for nonlinear systems may emerge which are not due to linearization but rather to the fact that the kinetic and transport parameters

[§]To whom correspondence should be addressed. E-mail: john.ross@stanford.edu.

© 2004 by The National Academy of Sciences of the USA

are close to each other, practically the same, for many labeled and unlabeled species, such a condition is referred to as the “neutrality condition.” In chemistry the neutrality condition is due to the kinetic isotope effect, which leads to generally small variations of the values of the rate coefficients; in population genetics it is because many mutations are neutral, that is, the demographic and transport parameters are practically the same for mutants and nonmutants. For neutral systems these linear response laws are exact even if the intrinsic dynamics is highly nonlinear. Another attractive feature is that both theories lead to simple physical interpretations for the susceptibility functions from the response laws; they are related to the probability density of the time necessary for crossing the system, which should simplify the analysis of experimental data.

The purpose of this article is to develop a unified linear response approach for nonlinear dynamics obeying neutrality conditions, valid both for continuous and discrete variables, time-dependent, and possibly space-dependent. We discuss the possible implications of this generalized approach for extracting mechanistic information from experimental data, with emphasis first on studies of gene expressions and comparison with linear approximations presented in the literature. Then we apply this theory to a nonlinear kinetic example for which a detailed analytical study is possible. We compare that exact solution with the first-order response theory based on appropriate tracer measurements and also compare it with the response of the linearized kinetic example. An important interest here is in the effects of error propagation in the analysis because of the application of measurements with poor precision. We conclude with some comments on the use of the theory presented here in improving several earlier approaches.

Outline of the Response Approach

We consider a complex system made up of different types of individuals (species), which can be atoms, molecules, quasi-particles, biological organisms, etc. The different types of individuals interact with each other and at the same time are involved in motion, which can be described by transport operators local in time. We denote by $\rho_u(\mathbf{r};t)$, $u = 1, 2, \dots$ the concentrations of the different species at position \mathbf{r} and time t , expressed in numbers of individuals per unit volume, and we assume that the rate of change of the species u , $R_u(t)$, can be expressed as a local, nonlinear function of the composition vector $\rho(\mathbf{r};t) = [\rho_u(\mathbf{r};t)]_{u=1,2,\dots}$ and of time t : $R_u(t) = R_u^+(\rho(\mathbf{r};t)) - R_u^-(\rho(\mathbf{r};t))$, $u = 1, 2, \dots$, where $R_u^+(\rho(\mathbf{r};t)) \geq 0$ and $R_u^-(\rho(\mathbf{r};t)) \geq 0$ are formation and consumption rates, respectively. The transport of the different species can be described by transport operators $\mathbb{L}_u \dots$, which are local in time and generally nonlocal in space. In this article we limit ourselves to transport operators of the “master” type:

$$\mathbb{L}_u \dots = \int_{\mathbf{r}'} [\dots W_u(\rho(\mathbf{r}';t), \mathbf{r}' \rightarrow \mathbf{r}) d\mathbf{r}' - \dots W_u(\rho(\mathbf{r};t), \mathbf{r} \rightarrow \mathbf{r}') d\mathbf{r}'], \quad [2]$$

where $W_u(\rho(\mathbf{r}';t), \mathbf{r}' \rightarrow \mathbf{r})$ are concentration-dependent transition rates. The operators describe the regular (Fick) and anti-crowding diffusion processes as particular cases. The evolution equations of the process are:

$$\partial \rho_u(\mathbf{r}, t) / \partial t = J_u^+(\mathbf{r}, t) - J_u^-(\rho(\mathbf{r}, t)) + R_u[\rho(\mathbf{r}, t)] + \mathbb{L}_u \rho_u(\mathbf{r}, t), \quad [3]$$

where $J_u^+(\mathbf{r}, t)$ is the input flux of species u , which is generally space- and time-dependent and can be controlled by the exper-

imenter, and $J_u^-(\rho(\mathbf{r}, t))$ is the output flux of species u , which is assumed to depend on the composition vector. Each species $u = 1, 2, \dots$ may exist in two different forms, “marked” and “not marked,” and both forms fulfill a “neutrality condition” (13); that is, their kinetic and transport properties are identical. In chemical kinetics a “marked species” can be a molecule containing a radioactive isotope and we neglect the kinetic isotope effect. In fluid mechanics a marked species can be a colored fluid for which the hydrodynamic properties (density, viscosity, and diffusion coefficients) are the same as the ones of the main fluid. In population genetics a marked species can be an individual carrying a neutral mutation, for which the main functions describing the vital statistics (natality and mortality functions and diffusion coefficients) are the same as for a nonmutant individual. In the following text we denote by $\rho_u(\mathbf{r}, t)$ and $\rho_u^*(\mathbf{r}, t)$, $u = 1, 2, \dots$, the concentrations of the not marked and marked species, respectively, and by $\rho_u^\Sigma(\mathbf{r}, t) = \rho_u(\mathbf{r}, t) + \rho_u^*(\mathbf{r}, t)$, $u = 1, 2, \dots$, the total concentrations of the species. At the beginning of the experiment the system contains only not marked species. The system need not be (but may be) in a stationary state. The experiment consists in varying the ratios,

$$\alpha_u(\mathbf{r}, t) = J_u^{+*}(\mathbf{r}, t) / J_u^{+\Sigma}(\mathbf{r}, t), \quad [4]$$

between the input fluxes of $J_u^{+*}(\mathbf{r}, t)$, the marked compounds and the total input fluxes, $J_u^{+\Sigma}(\mathbf{r}, t)$, with the preservation of the total input fluxes, $J_u^{+\Sigma}(\mathbf{r}, t)$. We record the response to these variations, the fractions,

$$\beta_u(\mathbf{r}, t) = J_u^{-*}[\rho(\mathbf{r}, t), \rho^*(\mathbf{r}, t)] / J_u^{-\Sigma}[\rho(\mathbf{r}, t), \rho^*(\mathbf{r}, t)], \quad [5]$$

of the marked outflow fluxes, $J_u^{-*}[\rho(\mathbf{r}, t), \rho^*(\mathbf{r}, t)]$, to the total output fluxes $J_u^{-\Sigma}[\rho(\mathbf{r}, t), \rho^*(\mathbf{r}, t)]$. We intend to obtain a relation between the excitation of the system, expressed by the fraction $\alpha_u(\mathbf{r}, t)$, and the response of the system, expressed by the fraction $\beta_u(\mathbf{r}, t)$.

We assume the existence of generalized neutrality conditions (13), in the form of scaling laws, which connect the kinetic and transport laws for the whole system to the corresponding laws for the marked and not marked species, respectively.

$$R_u^{+*}(\rho^*(\mathbf{r};t), \rho(\mathbf{r};t)) = \frac{\rho_u^*(\mathbf{r};t)}{\rho_u^*(\mathbf{r};t) + \rho_u(\mathbf{r};t)} R_u^+(\rho^*(\mathbf{r};t) + \rho(\mathbf{r};t)), \quad u = 1, 2, \dots \quad [6]$$

Eq. 6 expresses the fact that the marked and not marked species contribute equally to the transport process. We assume that the output fluxes, $J_u^{-\Sigma}[\rho(\mathbf{r}, t)]$, are expressed by kinetic laws, and thus they also obey a scaling condition similar to Eq. 6:

$$J_u^{-*}(\rho^*(\mathbf{r};t), \rho(\mathbf{r};t)) = \frac{\rho_u^*(\mathbf{r};t)}{\rho_u^*(\mathbf{r};t) + \rho_u(\mathbf{r};t)} J_u^{-\Sigma}(\rho^*(\mathbf{r};t) + \rho(\mathbf{r};t)). \quad [7]$$

We also introduce similar scaling conditions for the transport rates $W_u(\rho(\mathbf{r}';t), \mathbf{r}' \rightarrow \mathbf{r}) d\mathbf{r}$, $u = 1, 2, \dots$

$$\begin{aligned} W_{u*}(\rho^*(\mathbf{r};t), \rho(\mathbf{r}';t), \mathbf{r}' \rightarrow \mathbf{r}) d\mathbf{r} \\ = \frac{\rho_u^*(\mathbf{r}';t)}{\rho_u^*(\mathbf{r}';t) + \rho_u(\mathbf{r}';t)} W_{u*}(\rho^*(\mathbf{r}';t) + \rho(\mathbf{r}';t), \mathbf{r}' \rightarrow \mathbf{r}) d\mathbf{r}, \quad u = 1, 2, \dots \end{aligned} \quad [8]$$

The response experiment suggested before (ref. 12 and references therein and ref. 13) can be described in terms of two sets of nonlinear evolution equations of the type (Eq. 3), one set for the labeled concentration vector $\rho^*(\mathbf{r};t)$, and the second set for the total concentration vector $\rho(\mathbf{r};t) + \rho^*(\mathbf{r};t)$, respec-

tively. The transport and reaction rates in these two sets of evolution equations are connected to each other by means of the neutrality conditions (Eqs. 6–8). Together with suitable initial and boundary conditions, these two sets of evolution equations determine the time and space dependence of the total concentrations of the different species and of the marked species. Despite their nonlinearity, these two sets of evolution equations lead to a linear response law, which relates the excitation functions, $\alpha_u(\mathbf{r};t)$, to the response functions, $\beta_u(\mathbf{r};t)$. By using the neutrality conditions (Eqs. 6–8) and the definitions (Eqs. 4 and 5) for the excitation and response functions, after lengthy algebraic manipulations, the two sets of nonlinear evolution equations lead to a set of linear integro-differential equations, which relate the responses $\beta_u(\mathbf{r};t)$ to the excitations $\alpha_u(\mathbf{r};t)$. By representing the solution of this set of linear equations in terms of Green functions, we come to a generalization of Eq. 1:

$$\beta_u(\mathbf{r},t) = \sum_{u'} \int_{-\infty}^t \int_{\mathbf{r}'} \chi_{uu'}(\mathbf{r}',t' \rightarrow \mathbf{r},t) \alpha_{u'}(\mathbf{r}',t') d\mathbf{r}' dt', \quad [9]$$

where the space- and time-dependent susceptibility function $\chi_{uu'}(\mathbf{r}',t' \rightarrow \mathbf{r},t)$ is nonnegative and obeys the normalization condition:

$$\sum_{u'} \int_{-\infty}^t \int_{\mathbf{r}'} \chi_{uu'}(\mathbf{r}',t' \rightarrow \mathbf{r},t) d\mathbf{r}' dt' = 1. \quad [10]$$

It is possible to show that the susceptibility function has a simple physical interpretation. It is related to the probability density,

$$\varphi_{u'}(\tau, \Delta \mathbf{r} | \mathbf{r}, u; t) d\tau d\Delta \mathbf{r}, \text{ with } \sum_{u'} \int_0^\infty \int_{\mathbf{r}'} \varphi_{u'}(\tau, \Delta \mathbf{r} | \mathbf{r}, u; t) d\tau d\mathbf{r} = 1, \quad [11]$$

that a species u , which leaves the system at time t and position \mathbf{r} , entered the system as the species u' spent a residence time in the system between τ and $\tau + d\tau$, with $\tau = t - t'$, and has a displacement vector between $\Delta \mathbf{r}$ and $\Delta \mathbf{r} + d\Delta \mathbf{r}$, with $\Delta \mathbf{r} = \mathbf{r} - \mathbf{r}'$. We have

$$\varphi_{u'}(\tau, \Delta \mathbf{r} | \mathbf{r}, u; t) = \chi_{uu'}(\mathbf{r} - \Delta \mathbf{r}, t - \tau \rightarrow \mathbf{r}, t). \quad [12]$$

According to Eq. 12 the physical significance of the linear response law (Eq. 9) is straightforward, it expresses the contribution to the output fraction of marked species entering the system in different forms u' , at different initial positions \mathbf{r}' , and at different initial times, $t' = t - \tau$. The weight function (susceptibility function) attached to various initial positions and times is the conditional probability density of these three random variables. The theoretical results presented here include as particular cases our previous approaches to linear response for nonlinear systems with neutrality conditions (ref. 12 and references therein and ref. 13). To save space the derivation of Eqs. 9–12 is not given here. Similar derivations, corresponding to some particular cases, are given in refs. 12 and 13.

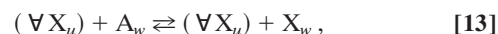
This article focuses on some possible applications of our approach for studying complicated chemical and biochemical systems. Our method makes it possible to use simple linear rules for exploring complicated nonlinear systems. A simple application is the study of connectivity among various chemical species in complicated reaction networks. In the simple case of homogeneous systems with time-invariant structure, the susceptibility

matrix $\chi = [\chi_{uu'}] = \chi(\tau)$ depends only on the transit time and not on time itself. Our theory shows that the matrix elements, $\chi_{uu'}(\tau)$, are proportional to the elements, $G_{uu'}(\tau)$, of a Green function matrix, $\mathbf{G}(\tau) = [G_{uu'}(\tau)]$, which is the exponential of a connectivity matrix \mathbf{K} , that is $\mathbf{G}(\tau) = \exp[\tau \mathbf{K}]$. It follows that, from a response experiment involving a system with time-invariant structure, it is possible to evaluate the connectivity matrix, \mathbf{K} , which contains information about the relations among the different chemical species involved in the reaction mechanism. Concerning the space variable \mathbf{r} from our approach, it can be used in different ways. If \mathbf{r} is a position variable in real space, then our approach can be used for analyzing response experiments in reaction-diffusion systems. Another possible application is to use our approach as a continuous formalism for the description of a complex discrete system; in this case, \mathbf{r} is not a real space position variable but a state vector describing the state of the system.

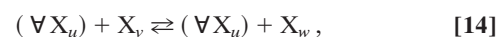
Error Analysis for a Test Model

To illustrate our approach, we consider a simple test model, which has the advantage that it can be studied analytically even for very large numbers of intermediates, making it suitable for the analysis of the interference between experimental errors and the errors due to linearization. This type of model, which is somewhat similar to Eigen's hypercycle model (15), was recently introduced by us in connection with a population genetic problem (M.O.V., L.L. Cavilli-Sforza, and J.R., unpublished data). The model used here is essentially a space-independent, homogeneous version of the model in M.O.V., L.L. Cavilli-Sforza, and J.R. (unpublished data). We assume that two types of chemical species are in the system: stable chemicals, A_ν , $\nu = 1, 2, \dots$, and active intermediates, X_u , $u = 1, 2, \dots$, and a very large supply of stable species A_ν , $\nu = 1, 2, \dots$ and their concentrations a_ν , $\nu = 1, 2, \dots$ are assumed to be constant and only the concentrations x_u , $u = 1, 2, \dots$ of the active intermediates are variable. We consider that the active intermediates replicate, transform into each other, and disappear through autocatalytic processes; moreover, we assume that all active intermediates have the same catalytic activity. Under these circumstances we can represent the chemical processes occurring in the system by the following reaction mechanism.

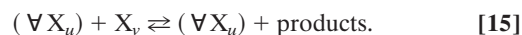
1. Replication processes



2. Transformation processes



3. Disappearance processes



In Eqs. 13–15 $(\forall X_u)$ denotes the pool of active intermediates. The kinetic evolution equations attached to the reactions (Eqs. 13–15) are

$$d\mathbf{x}/dt = \mathbf{x}_\Sigma \mathbf{A} \mathbf{x}, \quad [16]$$

where \mathbf{x} is the composition vector of the system, \mathbf{A} is a kinetic matrix that can be expressed in terms of the rate coefficients of the process, and $x_\Sigma = \sum_u x_u$ is the total concentration of active intermediates. Although the number of variables for this type of system can be very large, the integration of kinetic equations (Eq. 16) can be reduced to two numerical quadratures. This outcome is made possible by a nonlinear transformation of state variables involving an intrinsic timescale, $\theta(t) = \sum_u \int_0^t x_u(t') dt'$; with this intrinsic timescale the kinetic equations

(Eq. 16) can be expressed in terms of a linear matrix differential equation:

$$dx/d\theta = Ax. \quad [17]$$

If the kinetic matrix is constant, Eq. 17 can be solved either analytically by using the Sylvester theorem or the Laplace transformation or numerically. Finally, the relation between the intrinsic timescale θ and the laboratory timescale t can be determined from $t = \int_0^\theta [\sum_i x_{ii}(\theta')]^{-1} d\theta'$. A constant kinetic matrix is sufficient for representing response experiments for which the excitations can be represented by step functions. For more complicated excitations the kinetic matrix is time-dependent. In this case, the method of intrinsic time can be extended by representing the general solution of Eq. 17 in terms of a Green matrix function G , which can be evaluated by repeated numerical integration of an equation of the type $dG/dt = AG$ with initial conditions displaced at different initial times.

For illustration we start by analyzing a very simple type of response experiment involving the evaluation of self-replication constants. We consider the perturbation of the self-replication of a species by an input flux, whereas the response to this perturbation is given by the variation of the disappearance rate of the species due to the input perturbation. We study two types of excitation: (i) an increase of the input flux according to a step function and (ii) an excitation of the neutral type, where the total input flux is kept constant, but the fraction of a labeled compound in the input flux is varied, also as a step function. We assume that the measured values of the variation of the output flux (disappearance rate) are subject to experimental errors. To emulate the two types of response experiments, we solve the evolution equations (Eq. 16) exactly for the two types of response and add exactly the same errors to the two results. We take these solutions as our “experimental” values. Further on, we take the responses with errors and evaluate the self-replication rate from them. For the first type of emulated response experiment, type *i*, we evaluate the rate coefficient from the linearized evolution equation. For the neutral response experiment, type *ii*, we evaluate the self-replication constant by deconvoluting the susceptibility function from our exact response law (Eq. 9). In both cases, we study the variation of the relative (percent) output error of the evaluated self-replication rate in terms of the relative (percent) values of the measurement errors and of the relative, percent variations of the input flux. For the experiments of type *i* the variation of the input flux is expressed as the ratio between the flux increase due to excitation and the total value of the flux after the excitation has occurred. For the experiments of type *ii* the total flux is kept constant and the variation is the fraction of labeled compound in the input flux after the excitation has occurred.

Figs. 1 and 2 show the error of the evaluated self-replication rate constant as a function of the experimental error and of the variation of the input flux for emulated experiments of types *i* and *ii*, respectively. In the emulated experiments of type *i* for which the evaluation of the rate constant is based on linearized kinetic equations, the error of the evaluated rate constant depends strongly on the variation of the input perturbation. The range of the final output error (−40%, +10%) is distorted in comparison with the range of the experimental error (−20%, +20%). For small values of the input perturbation, between 20% and 40% the output error is surprisingly small, between 10% and 0%. As the input perturbation increases, the accuracy of the method deteriorates rapidly, and for large perturbations the output error is almost twice as big as the experimental error. For the emulated experiments of type *ii*, where the rate coefficient is evaluated from our exact response law (Eq. 9) without linearization, the situation is different. For input perturbations between 20% and 70% the error of the evaluated rate coefficient

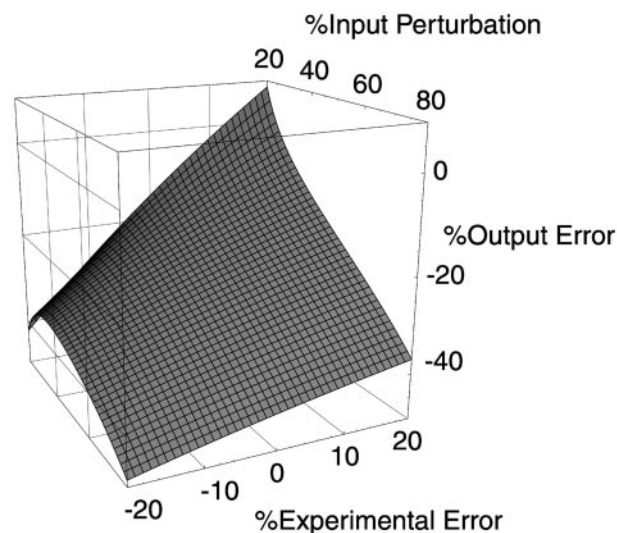


Fig. 1. The error of the evaluated self-replication rate constant (output error) versus the experimental error and the input perturbation for an emulated response experiment of type *i*. The rate constant is evaluated by using a linearized evolution equation. The range of the output error (−40%, +10%) is strongly distorted in comparison with the range of the experimental error (−20%, +20%). The output error varies a lot with the perturbation size; for large perturbations it is about twice the experimental error, whereas in other regions error compensation occurs.

has about the same range of variation as the experimental error (−20%, +20%) and does not depend much on the size of the perturbation. For very large input perturbations, between 70% and 80%, the output error increases abruptly. In Fig. 3 we show the difference of errors of the evaluated self-replication rate constant (output error) evaluated from emulated experiments of type *i*, with linearization, and type *ii*, without linearization, respectively. We notice that the two approaches lead to different

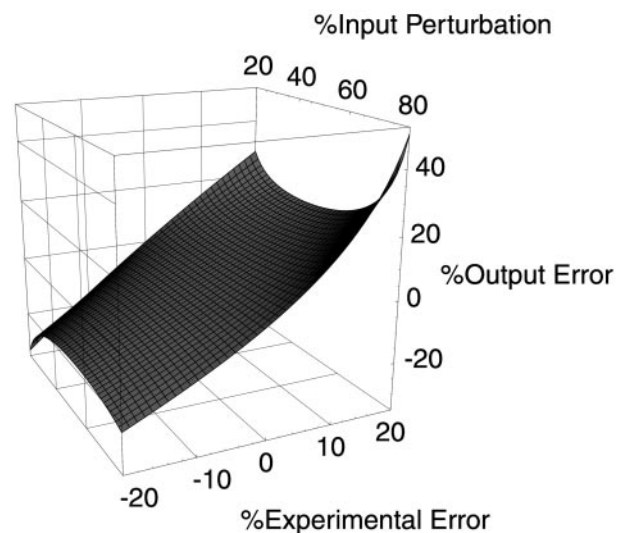


Fig. 2. The error of the evaluated self-replication rate constant (output error) versus the experimental error and the input perturbation for an emulated response experiment of type *ii*. The rate constant is evaluated from the susceptibility function computed from the exact response law (Eq. 9). The range of the output error is about the same as the range of the input error for input perturbations between 20% and 70% and does not depend much on the size of the perturbation. For input perturbations between 70% and 80%, the output error is bigger.

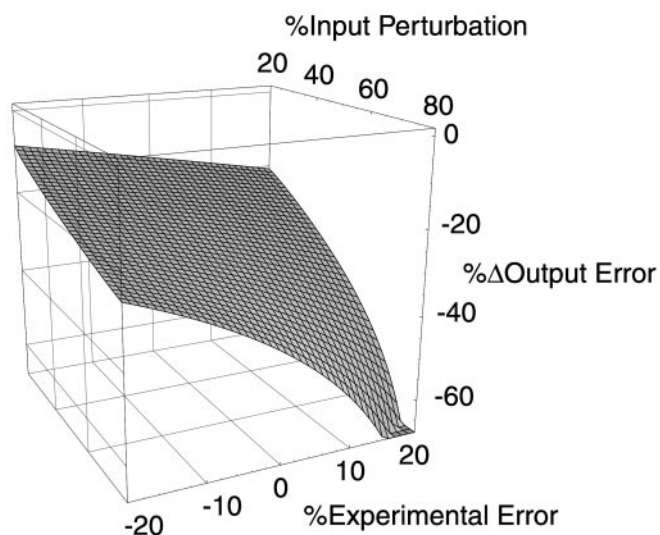


Fig. 3. The difference of errors of the evaluated self-replication rate constant (output error) evaluated from emulated experiments of types *i*, with linearization, and *ii*, without linearization, respectively. The figure shows that the variations of the input perturbation and of the experimental error have a different effect on the two types of response methods. The biggest difference occurs for large perturbations, because for large perturbations the linear approach is very inaccurate.

results; the evaluated values of the rate coefficient are different even for small experimental errors and input perturbations. The biggest differences occur for large perturbations, because for large perturbations the linear approach is very inaccurate.

The physical interpretation of the results presented in Figs. 1–3 follows. Like any autocatalytic processes, chemical reactions Eqs. 13–15 lead to saturation effects due to the balance between self-replication and consumption (disappearance) processes. The saturation effects are nonlinear and, as a result, the experimental errors propagate nonlinearly, which explains the error distortion displayed in Fig. 1. Even though the analysis of the process is based on linear equations, the intrinsic dynamics of the process is nonlinear and the errors are distorted. This error distortion interferes with the numerical errors because of linearization, resulting in Fig. 1. In some regions, these two errors can have opposite signs, resulting in error compensations, whereas in other regions the two errors add up, leading to large output errors. For neutral response experiments, type *ii* linearization is not necessary for the evaluation of the rate coefficients and the most important source of errors is the experimental errors. This is the reason why the output error follows closely the input error and almost no error compensation or error amplification occurs. Numerical errors, however, do have some influence in neutral experiments, especially for very-large-input perturbations. If the input flux contains a very large fraction of marked (labeled) compound the dynamics of the marker is close to saturation, and, in this area, small variations of the input flux lead to relatively large variations in the output, which produces numerical errors in the deconvolution of Eq. 9. This effect explains the spike in Fig. 2 for large perturbations. It follows that, even for response experiments of type *ii*, it is not recommended to use very-large-input perturbations. Nevertheless, it seems that the admissible input perturbations that produce reasonable

results are much larger than the admissible perturbations for experiments of type *i* for which the analysis is based on linearized evolution equations.

We have also emulated experiments that involve the measurement of more than one flux, to determine the cumulative effect of errors produced by more than one measurement. We have considered the evaluation of two transformation constants. Unfortunately, in this case, the results cannot be easily represented as compact, 3D graphs. The analysis revealed a complication typical for multiple flux measurements: experimental errors may lead to an apparent violation of the law of mass conservation. We considered only errors that preserve the mass conservation. We noticed the same qualitative features as in the self-replication constant. For experiments of type *i*, with linearization, in general, a distortion of the range of the output error exists compared with the range of the input error. We have also noticed error compensation and error accumulation; in some cases, they exist in more than one region and alternate, one compensation region followed by an accumulation region. Sometimes the amplification error tends to be very large; we have noticed output ranges of the error four to five times larger than the maximum range of the measurement errors of the fluxes. The emulated experiments of type *ii* are also affected by the cumulative measurement errors in a complicated way. A region exists where the range of the output error is about the same as the maximum range of the input experimental errors; however, the maximum admissible input fluxes are smaller than in a single-response experiment; in our simulations it varied from ~60–65% (compared with 80% for single-response experiments) to 35–40%.

Although the detailed results of our computations are probably model-specific we believe that the qualitative features discussed before have general significance. It is likely that the analysis of a response experiment based on linearized equations led to error compensation in some regions and to error amplification in other regions. The nonlinear saturation effects occurring in our model are likely to exist even though the autocatalytic steps are missing. For example, enzymatic reactions display saturation effects due to the limited supplies of the enzymes. The error compensation is as harmful as error amplification because, in general, it is impossible to predict *a priori* where it occurs. Concerning the neutral response experiments for which the linearization results are not necessary, it is likely that they lead to better results than the experiments of type *i* for which the analysis is based on the linearization of the kinetic equations. However, even for neutral experiments the use of very-large-input perturbations requires caution, because of the proximity to saturation, which results in numerical errors. Multiple measurements with large errors introduce further complications, which reduce the potential value of experiments of type *ii*, without linearization, for large perturbations.

In conclusion, the linearization of the evolution equations for the analysis of chemical and biochemical networks is unpredictably limited. For the linearization to be valid it is necessary to use small perturbations, for which the experimental errors are very large. In the articles that have appeared on this subject insufficient (or no) attention has been given to error accumulation and propagation (4–7). The response approaches developed here avoid linearization and, hence, are to be preferred.

This research was supported in part by the National Science Foundation.

1. Kohane, I. S., Kho, A. T. & Butte, A. J. (2003) *Microarrays for an Integrative Genomics* (MIT Press, Cambridge, MA).
2. Arkin, A., Shen, P. & Ross, J. (1997) *Science* **277**, 1275–1279.
3. Torralba, A., Yu, K., Peidong, S., Oefner, J. P. & Ross, J. (2003) *Proc. Natl. Acad. Sci. USA* **100**, 1494–1498.

4. Wang, W. J., Cherry, J. M., Botstein, D. & Li, H. (2002) *Proc. Natl. Acad. Sci. USA* **99**, 16893–16898.
5. Yeung, M. K. S., Tegner, J., Collins, J. J. (2002) *Proc. Natl. Acad. Sci. USA* **99**, 6163–6168.
6. Bruggemann, F. J., Westerhoff, H. V., Hoeck, J. B. & Kholodenko,

- B. N. (2002) *J. Theor. Biol.* **218**, 507–520.
7. Segre, D., Vitkup, D. & Church, G. M. (2002) *Proc. Natl. Acad. Sci. USA* **99**, 15112–15117.
8. D'Haeseleer, P., Liang, S. & Simogyi, R. (2000) *Bioinformatics* **16**, 707–726.
9. Dattagupta, S. (1987) *Relaxation Phenomena in Condensed Matter Physics* (Academic Press, New York).
10. Brezonik, P. L. (1994) *Chemical Kinetics and Process Dynamics in Aquatic Systems* (Lewis Publishers, Boca Raton, FL).
11. Edmonds, C. A., Lillie, A. S. & Cavalli-Sforza, L. L. (2004) *Proc. Natl. Acad. Sci. USA* **101**, 975–979.
12. Vlad, M. O., Ross, J., Moran, F. & Joel, Y. (2002) *Int. J. Bif. Chaos* **12**, 2599–2618.
13. Vlad, M. O., Moran, F., Tsuchiya, M., Cavalli-Sforza, L. L., Oefner, P. J. & Ross, J. (2002) *Phys. Rev. E* **65**, 061110-1–061110-17.
14. Vlad, M. O. & Ross, J. (2002) *Phys. Rev. E* **66**, 061908-1–061908-14.
15. Eigen, M. (1979) *The Hypercycle: A Principle of Natural Self Organization* (Springer, Berlin).